Doubly Robust Estimators with Weak Overlap

Yukun Ma
U Rochester

Pedro H.C. Sant'Anna
Emory University

Yuya Sasaki
Vanderbilt U

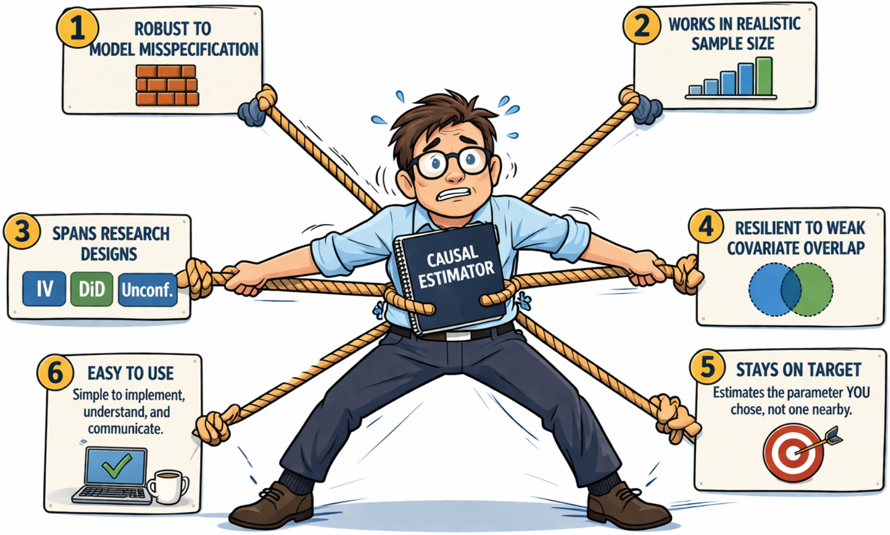
Takuya Ura
UC Davis

University of Georgia
April 2026

A Desiderate for Causal Inference Procedures

What applied researchers want from a causal estimation procedure:

1. **Robust to model misspecification** — we all make mistakes
2. **Works in realistic sample sizes** — not only with very large admin data
3. **Spans research designs** — unconfoundedness, IV / LATE, DiD, ...
4. **Resilient to weak covariate overlap** — the tension with rich conditioning sets
5. **Stays on target** — it estimates the parameter <u>you</u> chose, not one nearby


1 ROBUST TO MODEL MISSPECIFICATION
2 WORKS IN REALISTIC SAMPLE SIZE
3 SPANS RESEARCH DESIGNS
IV
DiD
Unconf.
4 RESILIENT TO WEAK COVARIATE OVERLAP
5 STAYS ON TARGET
Estimates the parameter YOU chose, not one nearby.
6 EASY TO USE
Simple to implement, understand, and communicate.
CAUSAL ESTIMATOR

All Models Are Wrong

> *All models are wrong, but some are useful.*
>
> --- George E.P. Box (1976)

- Motivates simple regression specifications that are considered "good enough:"

$$Y = \alpha + \tau D + X'\theta + u$$

- τ is usually interpreted as "the" causal effect of interest.
- Treatment effects heterogeneity $\rightarrow$ these linear regressions are misspecified.

What kind of causal parameter τ actually recover?

The Reverse Engineering Approach

$\rightarrow$ **Reverse engineering**: run a regression first, then figure out what it estimates.

- **Under unconfoundedness:** $\tau_{\text{OLS}} \neq \text{ATE, ATT, ATU}$ (Słoczyński, 2022)
- **Under IV / LATE:** $\tau_{\text{2SLS}} \neq \text{LATE}$ (Słoczyński, 2026; Blandhol, Bonney, Mogstad and Torgovitsky, 2026)
- **Under parallel trends / DiD:** $\tau_{\text{TWFE}} \neq \text{ATT}$ (Caetano and Callaway, 2024)

More details

- **OLS under unconfoundedness** (Słoczyński, 2022):
A convex combination of ATT and ATU — but the **weights are reversed**: the more units are treated, the **less** weight OLS places on ATT.

- **2SLS under IV / LATE** (Słoczyński, 2026; Blandhol et al., 2026):
Linear IV can place **negative weights** on some conditional LATEs — the estimand can be negative even when every individual effect is positive.

- **TWFE under parallel trends** (Caetano and Callaway, 2024):
Even with two periods, TWFE can implicitly impose a **different PT assumption** than the one you thought you imposed; negative-weight problems appear, too.

My concern with the Reverse Engineering Approach

> ``*When I'm not near the parameter I love, I love the parameter I'm near.*''
>
> *--- Art Goldberger, paraphrasing the song When I'm Not Near the Girl I Love (Finian's Rainbow, 1947)*

Forward-Engineering Approach

→ **Forward engineering**:

1. Starts with the causal question
2. Make clear identification and statistical assumptions
3. Build an estimator that targets your parameter of interest.
4. **Bonus:** Enjoy some attractive statistical guarantees.

Forward-Engineering Approach

→ **Forward engineering**:

1. Starts with the causal question
2. Make clear identification and statistical assumptions
3. Build an estimator that targets your parameter of interest.
4. **Bonus:** Enjoy some attractive statistical guarantees.

> **Doubly Robust (DR) estimators are particularly well-suited here**

Forward-Engineering Approach via DR

	PS correct	PS wrong
RA correct	✓ Consistent	✓ Consistent
RA wrong	✓ Consistent	× Inconsistent

The ATE Case

- Let $p_D(X) = \mathbb{E}[D{=}1 \mid X]$ (propensity score) and $m_Y(d, X) = \mathbb{E}[Y \mid D{=}d, X]$ (outcome regression).

- The DR/AIPW estimand (Robins, Rotnitzky and Zhao, 1994; Słoczyński and Wooldridge, 2018):

$$\text{ATE}_{\text{DR}} = \mathbb{E}\Big[m_Y(1,X) - m_Y(0,X) + w_1^{\text{ATE}}\big(Y - m_Y(1,X)\big) - w_0^{\text{ATE}}\big(Y - m_Y(0,X)\big)\Big]$$

with Hájek-type normalized weights

$$w_1^{\text{ATE}} = \frac{D \,/\, p_D(X)}{\mathbb{E}\big[D \,/\, p_D(X)\big]}, \qquad w_0^{\text{ATE}} = \frac{(1-D) \,/\, (1-p_D(X))}{\mathbb{E}\big[(1-D) \,/\, (1-p_D(X))\big]}$$

The LATE Case

- Let $p_Z(X) = \mathbb{E}[Z \mid X]$ (instrument PS), $m_R^{\text{LATE}}(z, X) = \mathbb{E}[R \mid Z{=}z, X]$ for $R \in \{Y, D\}$.
- The DR Wald ratio (Belloni, Chernozhukov, Fernández-Val and Hansen, 2017; Słoczyński, Uysal and Wooldridge, 2025):

$$\text{LATE}_{\text{DR}} = \frac{\mathbb{E}\left[\Delta m_Y^{\text{LATE}}(X) + w_{Z=1}^{\text{LATE}}\left(Y{-}m_Y^{\text{LATE}}(1, X)\right) - w_{Z=0}^{\text{LATE}}\left(Y{-}m_Y^{\text{LATE}}(0, X)\right)\right]}{\mathbb{E}\left[\Delta m_D^{\text{LATE}}(X) + w_{Z=1}^{\text{LATE}}\left(D{-}m_D^{\text{LATE}}(1, X)\right) - w_{Z=0}^{\text{LATE}}\left(D{-}m_D^{\text{LATE}}(0, X)\right)\right]}$$

where $\Delta m_R^{\text{LATE}}(X) = m_R^{\text{LATE}}(1, X) - m_R^{\text{LATE}}(0, X)$ and normalized weights

$$w_{Z=1}^{\text{LATE}} = \frac{Z \,/\, p_Z(X)}{\mathbb{E}\left[Z \,/\, p_Z(X)\right]}, \qquad w_{Z=0}^{\text{LATE}} = \frac{(1{-}Z) \,/\, (1{-}p_Z(X))}{\mathbb{E}\left[(1{-}Z) \,/\, (1{-}p_Z(X))\right]}$$

The DiD Case

- Let G be the period of first treatment, $C_{g,t}$ the comparison indicator, $p_{g,t}(X) = \mathbf{Pr}[G = g \mid X, G + C_{g,t}{=}1]$ the generalized PS, $m_{g,t}(X) = \mathbb{E}[Y_t - Y_{g-\delta-1} \mid X, C_{g,t}{=}1]$, and $\delta \geq 0$ anticipation periods.

- The DR estimand for the group-time ATT: (Callaway and Sant'Anna, 2021)

$$\mathrm{ATT}_{\mathrm{DR}}(g,t) = \mathbb{E}\left[\left(w^{\mathrm{DiD}}_{G=g}(G) - w^{\mathrm{DiD}}_{g,t}(G,X)\right)\left(Y_t - Y_{g-\delta-1} - m_{g,t}(X)\right)\right]$$

with weights

$$w^{\mathrm{DiD}}_{G=g}(G) = \frac{\mathbb{1}\{G{=}g\}}{\mathbb{E}[\mathbb{1}\{G{=}g\}]}, \quad w^{\mathrm{DiD}}_{g,t}(G,X) = \frac{C_{g,t}\, p_{g,t}(X) \,/\, (1{-}p_{g,t}(X))}{\mathbb{E}\left[C_{g,t}\, p_{g,t}(X) \,/\, (1{-}p_{g,t}(X))\right]}$$

DR estimators sound great.

Where is the catch?

The Overlap Problem

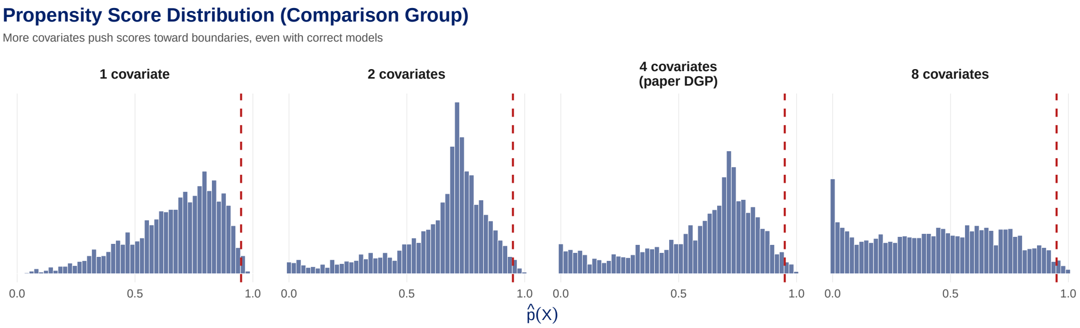
The Covariate Tension

Rich covariates make identification credible...

- Interviewer fixed effects for the test score gap
- County characteristics for parallel trends in DiD
- Income controls for instrument validity in IV

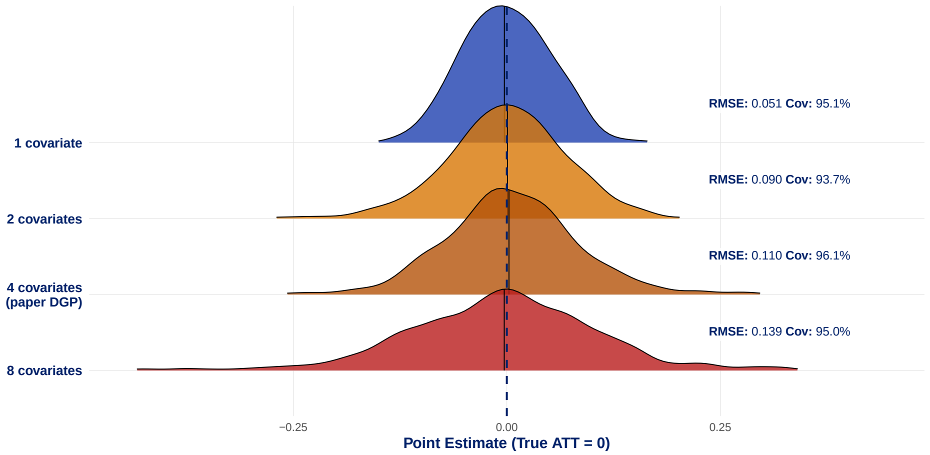
...but the same covariates that make identification **credible** can make estimation and inference **challenging** through weak overlap (D'Amour, Ding, Feller, Lei and Sekhon, 2021).

Usually, More Covariates → Worse Overlap



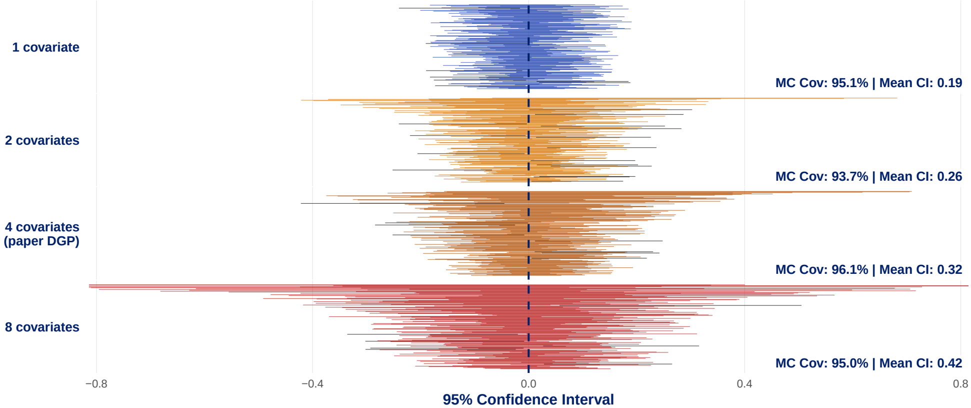
DiD simulation, $n = 5{,}000$, Student-$t(10)$ covariates. All models are correctly specified.

Even With Correct Models, Estimates Get Noisy



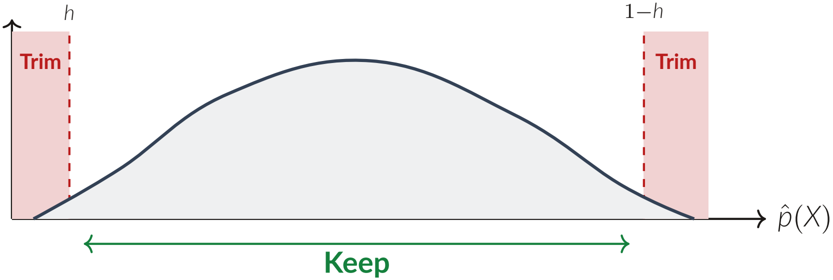
Each panel adds covariates to the PS model. RMSE nearly triples from 1 to 8 covariates.

Inference Becomes Less Informative



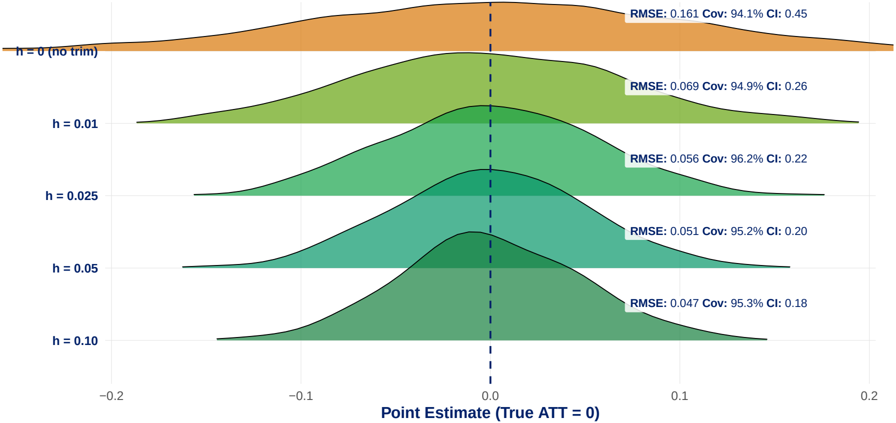
Black segments = non-coverage. CIs more than double from 1 to 8 covariates.

The Natural Fix: Trim the Extremes



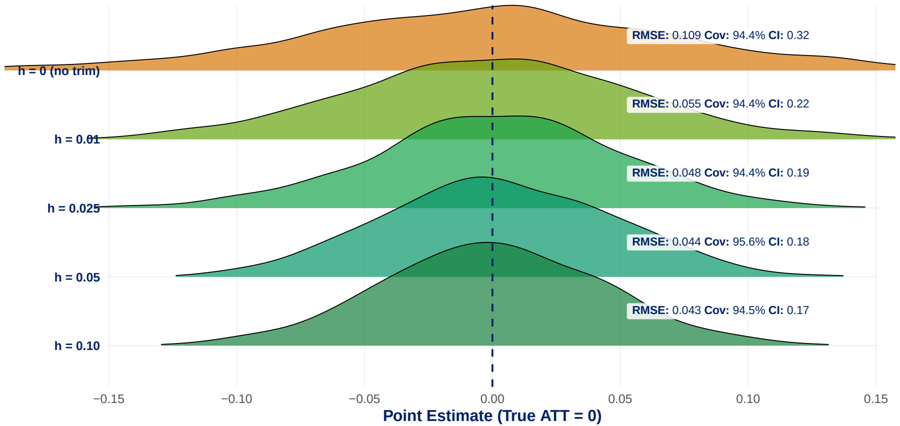
Trim obs with extreme PS. Variance drops immediately ✓

Trimming Comparison Units: All Models Correct - 8X



Same DGP, 8 covariates. Trimming cuts RMSE by $2.9\times$ and CI width by $2.4\times$.

Trimming Comparison Units: All Models Correct, 4X

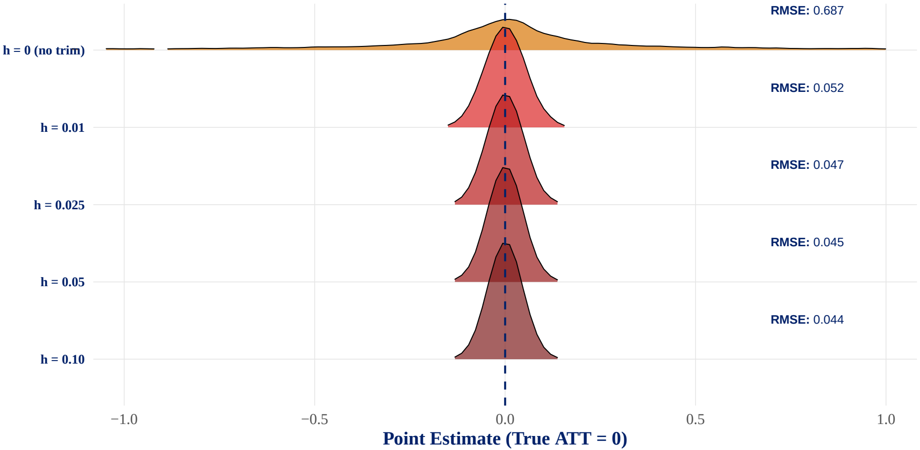


Paper DGP with 4 covariates. Both models correctly specified. Trimming reduces RMSE and tightens CIs.

Are we done?

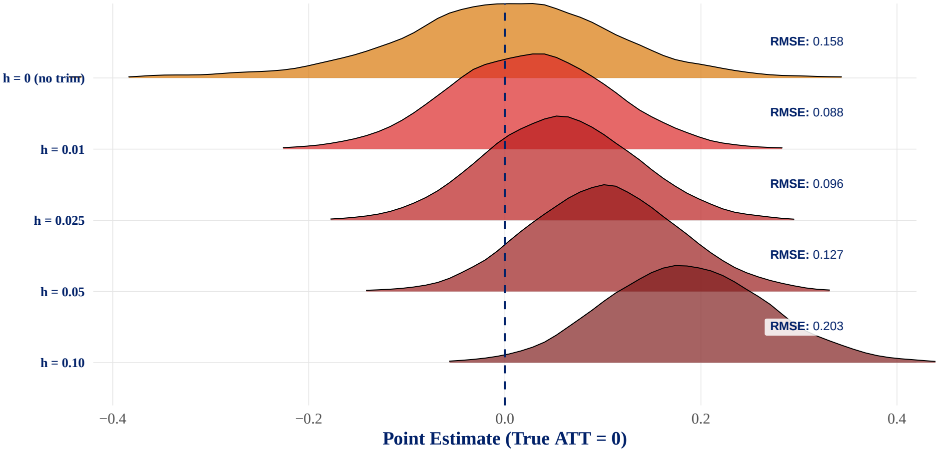
What if
some of the models
are misspecified?

Trimming + DR: RA Correct, PS Misspecified



PS misspecified, outcome regression correct. Nominal 95% coverage: **achieved** ✓

Trimming + DR: PS Correct, RA Misspecified



Nominal 95% → actual coverage collapses to 34%.

Trimming

destroys

double robustness

Can't estimate your parameter?

Just change the question!


10

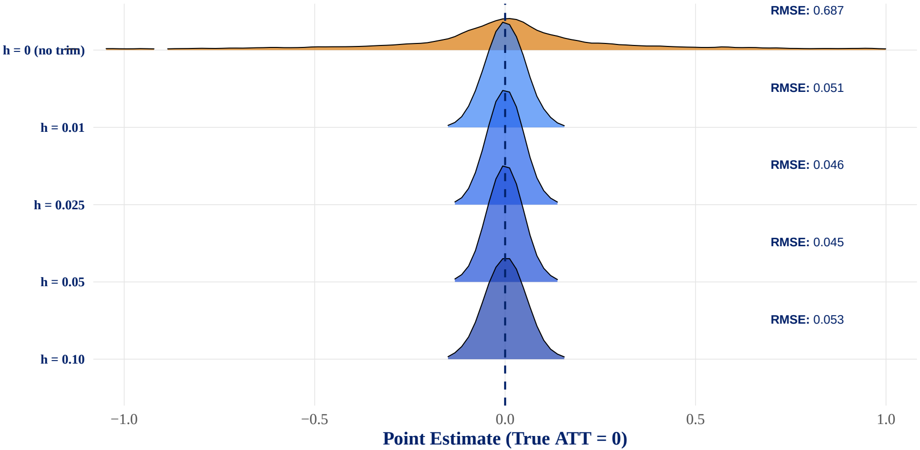
Keep your estimand.

Handle weak overlap.

Don’t move the goalposts — fix the kicker.

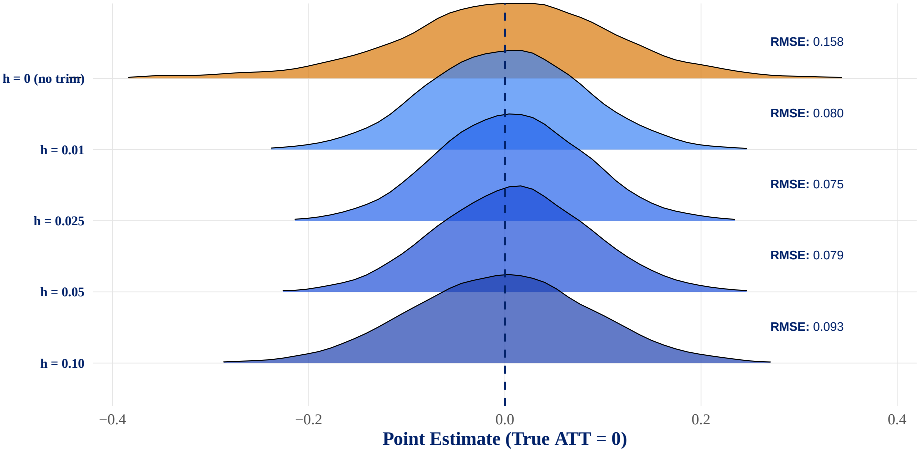

10

DR-BC: PS Misspecified, RA Correct



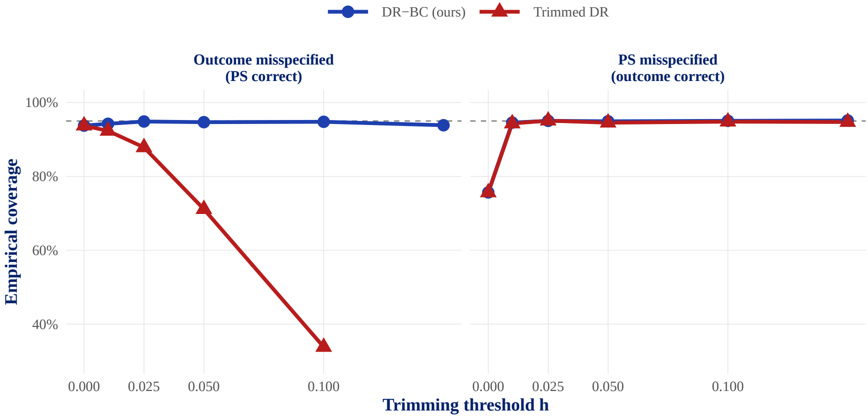
DR-BC preserves double robustness. Coverage ≈ 95% across all trimming thresholds ✓

DR-BC: RA Misspecified, PS Correct



The route that trimming destroyed — DR-BC **restores it**. Coverage 94–95% ✓

Coverage as a Function of h



DR-BC: stable near 95% across all h. Trimmed DR: collapses when PS is the correct model.

Why Trimming Breaks DR and How We Fix It

DR Estimands Share a Common Structure

- All DR estimands — ATE, LATE, DiD ATT — are functions of **ratio moments**:

$$\alpha_\ell = \mathbb{E}\left[\frac{B_\ell}{A_\ell}\right], \qquad \ell = 1, \ldots, L$$

  where A_ℓ involves the PS (can be near zero) and B_ℓ contains a group indicator $\times$ outcome residual

- **Example (ATE):** $A = p(X),\ B = D \cdot (Y - m(1, X)) \ \Rightarrow\ B/A$ is the IPW residual

- Trimming discards observations where $|A_\ell| < h$. This **permanently changes the estimand**:

$$\mathbb{E}\left[\frac{B_\ell}{A_\ell} \cdot \mathbb{1}|A_\ell| \geq h\right] \neq \mathbb{E}\left[\frac{B_\ell}{A_\ell}\right]$$

  The gap does not vanish for fixed h

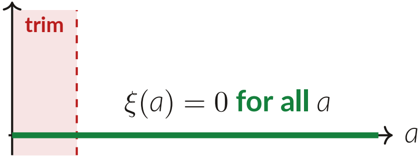
Why Trimming Breaks One Route But Not the Other

Define the **conditional mean**: $\xi(a) = \mathbb{E}[B \mid A = a]$

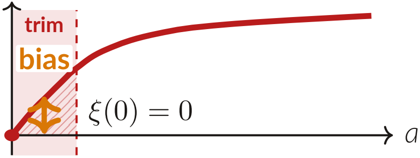
- **Outcome regression correct:**
  - DR residual $Y - m(X)$ has $\mathbb{E}[Y - m(X) \mid X] = 0$
  - Since A is a function of X, by iterated expectations: $\xi(a) = 0$ for **all** a
  - Trimming removes observations contributing zero $\to$ **no bias**

- **Propensity score correct:**
  - At $A = 0$, the PS is at its boundary $\to$ one group has probability zero
  - The group indicator in B forces $B = 0$ a.s. $\to \xi(0) = 0$
  - But $\xi(a) \neq 0$ for $a > 0$ in general
  - Trimming discards observations with $\xi \neq 0 \to$ **bias!**

The Asymmetry, Visually



Outcome correct: bias $= 0$ ✓



PS correct: bias $\neq 0$ ×

> Trimming is safe exactly when the outcome model is correct — <u>precisely when you don't need it!</u>

But notice ...

Under either DR route, $\xi(0) = 0$.

$$\underbrace{\frac{\xi(a)}{a}}_{\text{“}0/0\text{”}} \xrightarrow[a \to 0]{} \xi'(0)$$

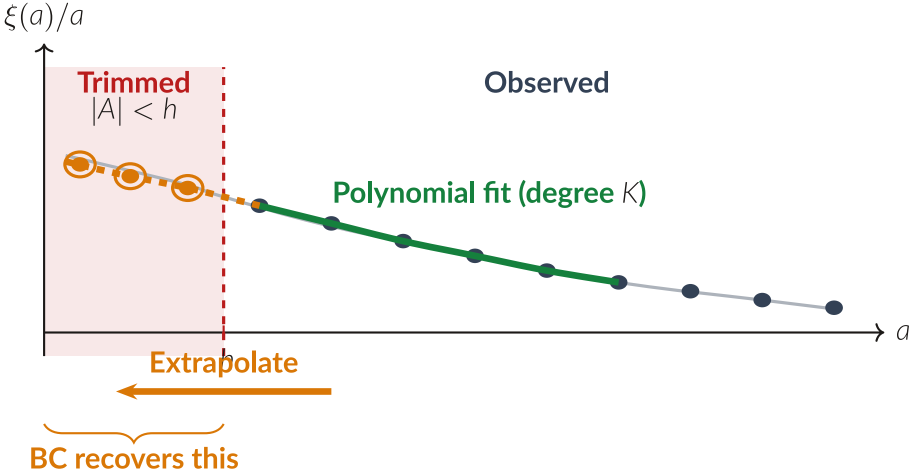
$B/A = \xi(A)/A$ has a **removable singularity** –
the “$0/0$” is a smooth function in disguise.

Polynomial Reconstruction

- We need $\mathbb{E}[B/A]$, but B/A is unstable near $A = 0$
- Rewrite: $B/A = \xi(A)/A$, which is a "$0/0$" form at the boundary
- By L'Hôpital's rule: $\xi(a)/a \to \xi'(0)$ as $a \to 0$
- So $\xi(a)/a$ is a **smooth function** near $a = 0$ — a low-degree polynomial can approximate it

$\Rightarrow$ **Fit a polynomial to $\xi(a)/a$ using kept observations,**
then extrapolate to reconstruct what trimming discarded

Polynomial Reconstruction



Hollow circles = reconstructed (not observed). Approximation error is $O(h^{k+1})$.

The DR-BC Estimator

- Replace each ratio moment with its **bias-corrected** version:

$$\hat{\alpha}_\ell(h) \;=\; \underbrace{\frac{1}{n}\sum_i \frac{B_{\ell,i}}{A_{\ell,i}}\mathbb{1}|A_{\ell,i}| \geq h}_{\text{Trimmed mean}} \;+\; \underbrace{\sum_{\kappa=1}^{k} \frac{\bar{A}_\ell^{\kappa-1}(h)}{\kappa!}\hat{\xi}_\ell^{(\kappa)}(0)}_{\text{Bias correction}}$$

where $\bar{A}_\ell^{\kappa-1}(h) = \frac{1}{n}\sum_i A_{\ell,i}^{\kappa-1}\mathbb{1}|A_{\ell,i}| < h$

- **One polynomial regression**: regress B_ℓ on a degree-K polynomial in $A_\ell \rightarrow$ evaluate $\hat{\xi}_\ell^{(\kappa)}(0)$
- Plug corrected $\hat{\alpha}_\ell$ into **any** DR formula: $\hat{\theta} = \Lambda(\hat{\alpha}_1, \ldots, \hat{\alpha}_L)$
- In practice: $h{=}0.05$, $k{=}1$ (first-order), $K{=}3$ (cubic) — a single tuning choice

The Key Conditions

- Decompose: $\hat{\theta} - \theta_0 = \underbrace{(\hat{\theta} - \theta_h)}_{\text{estimation}} + \underbrace{(\theta_h - \theta_0)}_{\text{trimming bias}}$
- **Approximate DR.** Under either route: $\theta_h(\gamma^*) = \theta_0 + o(h^k)$
- **Boundary condition.** $\xi_\ell(0; \gamma^*) = 0$ under either DR route
  - PS correct: group indicator forces $B_\ell = 0$ at boundary
  - Outcome correct: $\xi_\ell(a) = 0$ for all a (residual moments)
- **Smoothness.** $\xi_\ell(\cdot)$ is $(k+1)$-times differentiable near 0; $\Lambda(\cdot)$ smooth

Full details →

The Influence Function

IF of each bias-corrected moment $\hat{\alpha}_\ell$ has **four components**:

$$\omega_\ell = \underbrace{\frac{B_\ell}{A_\ell}\mathbb{1}\{|A_\ell| \geq h\}}_{\text{trimmed ratio}} + \underbrace{\sum_{\kappa=1}^{k} \frac{A_\ell^{\kappa-1}\mathbb{1}\{|A_\ell|<h\}}{\kappa!}\,\xi_\ell^{(\kappa)}(0)}_{\text{bias correction}} + \underbrace{\text{sieve IF}}_{\psi_{\ell,\kappa}} + \underbrace{\text{first-stage IF}}_{\phi_\ell}$$

- Overall IF via the delta method: $\varphi = \sum_{\ell=1}^{L} \Lambda_\ell \cdot \omega_\ell$
- Key difficulty: $\theta_0 = \Lambda(\alpha_1, \ldots, \alpha_L)$ is a **nonlinear function of multiple ratio moments** with heterogeneous convergence rates

Full details →

Main Result

Theorem. Under Assumptions 1–4:

(i) $\hat{\theta} - \theta_0 = (\mathbb{E}_n[\cdot] - \mathbb{E}[\cdot])[\varphi] + o_p(n^{-1/2})$

(ii) If in addition $\mathbb{E}[\varphi^2]$ is bounded away from zero and a Lyapunov condition holds, then $(\hat{\theta} - \theta_0) \Big/ \sqrt{\mathbb{E}[(\varphi - \mathbb{E}[\varphi])^2]/n} \xrightarrow{d} \mathcal{N}(0, 1)$

✓ **Analytic SEs** from the closed-form IF — no bootstrap; clustering + bands straightforward

✓ Rate conditions: $nh^4 \to \infty$ (lower) and $nh^{2k} = O(1)$ (upper); variance allowed to diverge

✓ Verified design-by-design for ATE, LATE, and DiD ATT

Applications

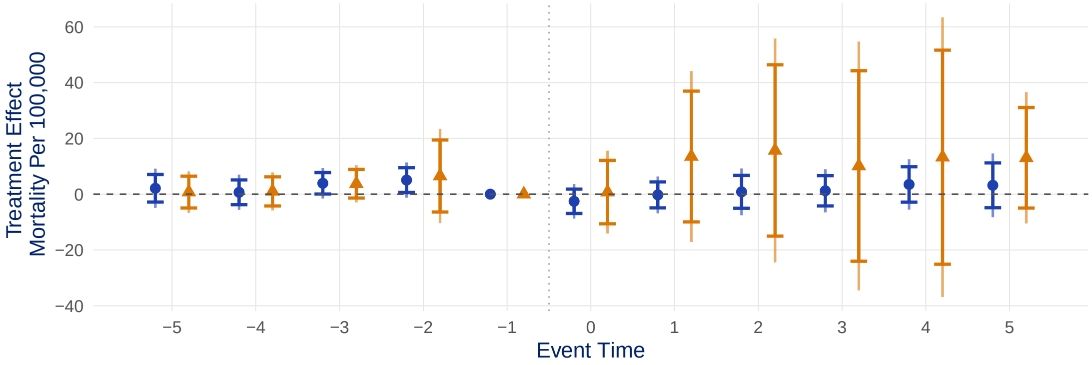
Four applications
three designs

one method

Medicaid Expansion & Mortality (DiD)

- **Stakes**: 22 states expanded Medicaid in 2014 under the ACA, 7 more by 2019 — did it save lives?
- **Data**: County-level all-cause mortality, ages 20–64, 2009–2019 (Baker, Callaway, Cunningham, Goodman-Bacon and Sant'Anna, 2025)
- **Design**: Staggered DiD (Callaway and Sant'Anna, 2021), never-treated as comparison
- **Covariates**: Poverty, unemployment, income, demographics (6 variables)
- **Problem**: Sharp covariate differences → PS near 1; up to 3.3% of comparison units trimmed — yet these carry extreme IPW weights

Medicaid: Event Study



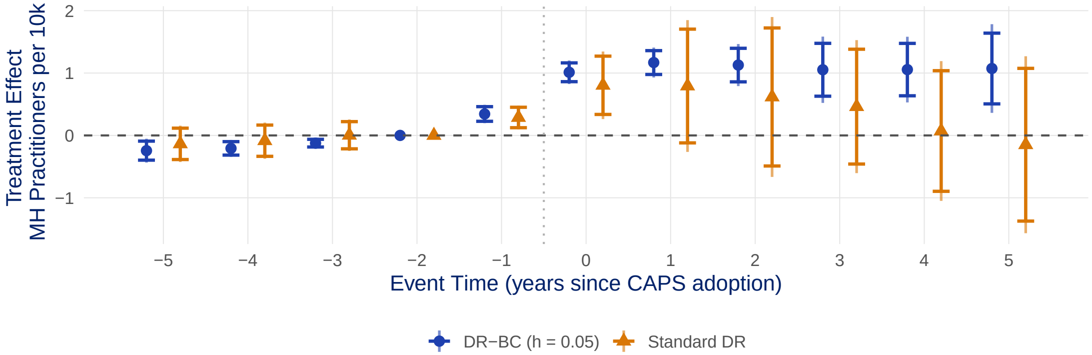
Std DR: simult. band $\sim$68 deaths/100K on avg (up to 101). DR-BC: $\sim$17 deaths/100K. Avg SE ratio: 3.5$\times$ (up to 4.6$\times$).

$3.5\times$ SE reduction $\approx 12\times$ more data for the same power.

Brazil Mental Health Reform (DiD)

- **Setting**: Brazil's 2002 psychiatric reform — staggered rollout of CAPS (community mental health centers) (Dias and Fontes, 2024)
- **Design**: DR DiD (Callaway and Sant'Anna, 2021) with 29 covariates + state FEs + population weights
- **Overlap challenge**: Rich covariates needed for parallel trends, but they limit overlap
- **Key test**: Does DR-BC matter even with very little trimming? (only 0.2% trimmed in the most affected cell)

Brazil: Event Study



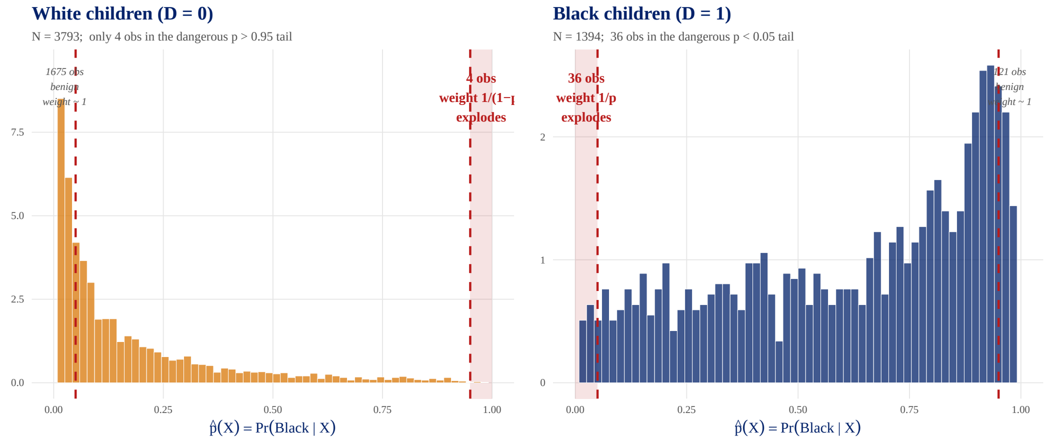
DR-BC: ~ 1 practitioner per 10K pop. post-adoption. **SE ratio: 2.8$\times$.**

Confirms DR-BC gains in a published AEJ:Policy setting with rich covariates.

Black–White Test Score Gap (ATE)

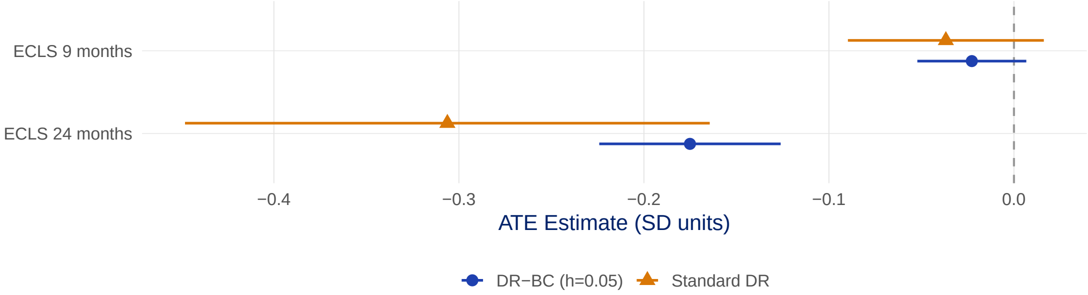
- **Question**: Racial test score gap at 9 and 24 months (Fryer and Levitt, 2013), treating race as the "treatment" (covariate-adjusted comparison)
- **Design**: DR for the ATE with the full Fryer–Levitt covariate set — family background, parental age polynomials, and **interviewer fixed effects**
- **The overlap problem**: Some interviewers tested almost only Black or almost only White children $\to \hat{p}(X) \to 0$ or $\to 1$
- **Only $\sim$0.7% of obs** (36–38 units) carry unstable weights, but 35–38% share a covariate cell with them — trimming operates cell-wise, so the BC reconstructs the whole region

Fryer–Levitt: Propensity Score Distribution



Each arm's dangerous tail is on a different side. Only ∼0.7% of obs carry unstable weights.

Fryer–Levitt: Results



DR-BC is more precise than OLS — while targeting the ATE.
SE ratio vs. Standard DR: 2.9$\times$ (24 months), 1.8$\times$ (9 months).

401(k) and Household Savings: Background

- **Question**: Does 401(k) participation increase household savings, or merely crowd out other forms of saving?

- **Literature**: A long-running debate
  - Abadie (2003): semiparametric IV/LATE — eligibility as instrument for participation
  - Benjamin (2003): propensity score subclassification — eligibility increases saving
  - Chernozhukov and Hansen (2004): instrumental quantile regression — positive effect across the distribution
  - Słoczyński et al. (2025): DR Wald ratio for the LATE — revisits Abadie's framework

- All approaches exploit employer-determined **eligibility** as an instrument for participation (employees do not choose whether their employer offers a 401(k) plan)

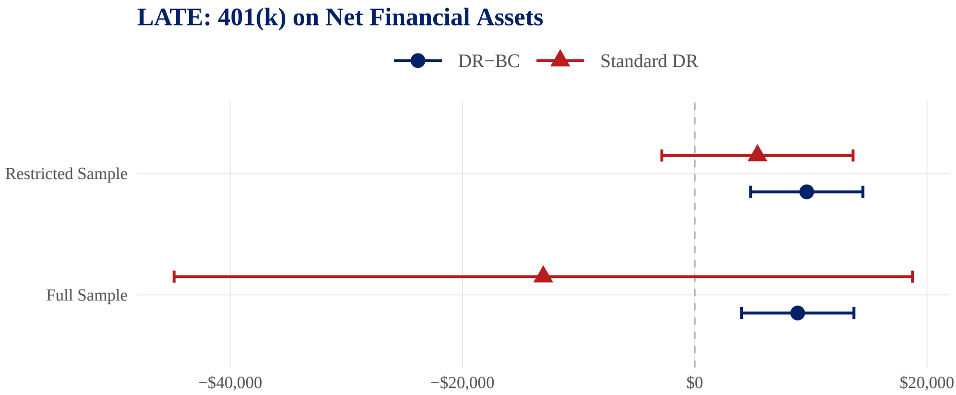
401(k): Data and the Overlap Problem

- **Data**: 1991 Survey of Income and Program Participation (SIPP)
- **Two samples**:
  - **Full sample**: positive income (Benjamin, 2003; Chernozhukov and Hansen, 2004), $n = 9{,}910$
  - **Restricted sample**: income $\in$ [\$10K, \$200K] (Abadie, 2003), $n = 9{,}275$
- **Why the restriction?** Abadie (2003, p. 249):

> " *``Outside this interval, 401(k) eligibility in the sample is rare."*

- **Our finding**: 4 of the 7 extreme-PS observations **survived** Abadie's income restriction — ad hoc sample restrictions may fail to address instability

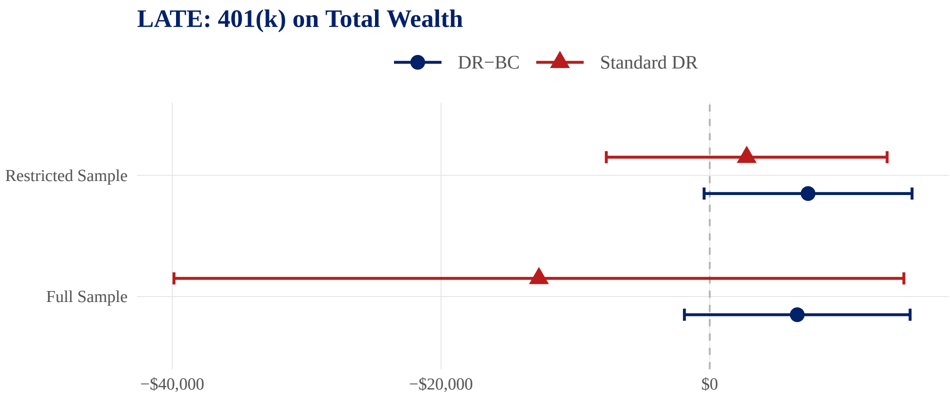
401(k): LATE on Net Financial Assets



Standard DR: $-\$13{,}042$ **(SE \$16,223) — wrong sign.** **DR-BC:** $+\$8{,}864$ **(SE \$2,471).**

SE ratio: 6.6$\times$ **(**$\approx 44\times$ **more data).**

401(k): LATE on Total Wealth



Standard DR: $-\$12{,}715$ **(SE \$13,859).** **DR-BC:** $+\$6{,}514$ **(SE \$4,286).** **SE ratio:** **3.2**$\times$.

Both outcomes positive and significant, consistent with Abadie (2003).

Conclusion

The Punchline

> SE reductions of 1.8–6.6× across four applications and three designs.
> Equivalent to 3–43× more data for the same power.

Summary of Results

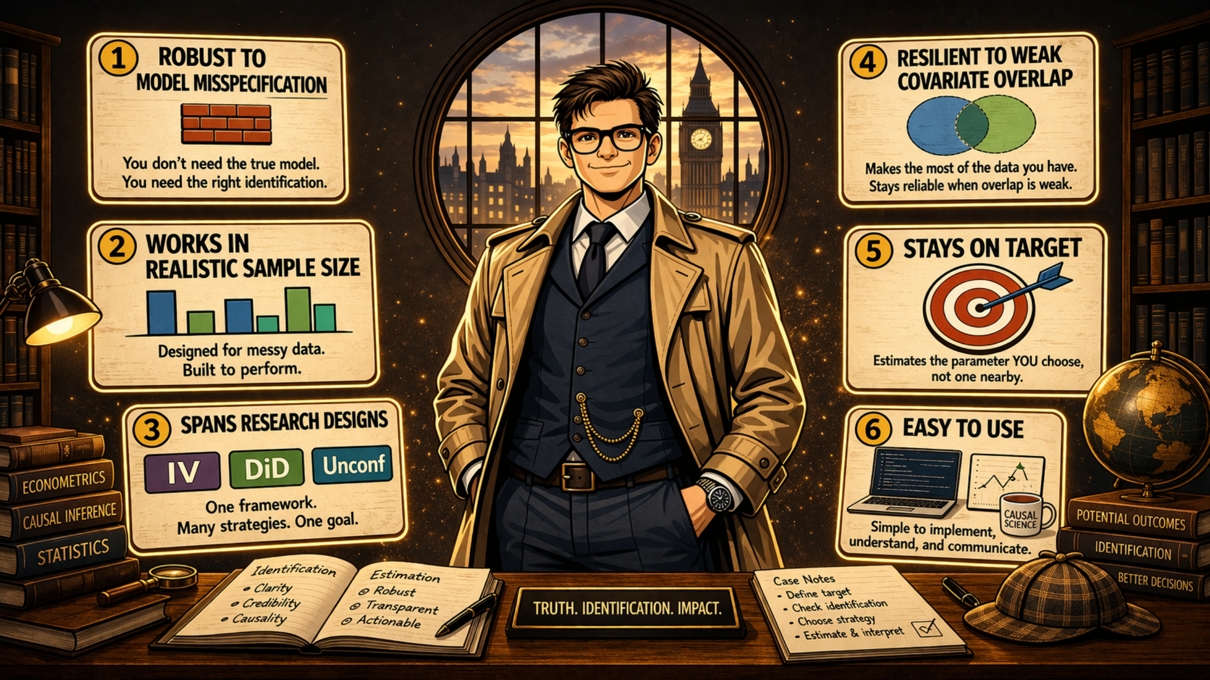
Application	Design	Trimmed	SE ratio	Data equiv. ($\approx c^2$)
Medicaid & mortality	DiD	up to 3.3%	**3.5**$\times$	**13**$\times$
Brazil mental health	DiD	up to 0.2%	**2.8**$\times$	**8**$\times$
Black–White test gap (24mo)	ATE	∼0.7%	**2.9**$\times$	**8**$\times$
401(k) savings (NFA)	LATE	2 of 9,910	**6.6**$\times$	**43**$\times$

A $c\times$ SE reduction $\approx c^2\times$ more data for the same power.

The Wishlist, Revisited

- ✓ **Robust to misspecification** — DR-BC preserves the DR guarantee
- ✓ **Works in realistic *n*** — applications at $n \approx 5\text{k}\text{–}10\text{k}$; theory + sims confirm
- ✓ **Spans research designs** — ATE, LATE, DiD ATT from one formula
- ✓ **Resilient to weak overlap** — $1.8\times$–$6.6\times$ SE gains across 4 applications
- ✓ **Stays on target** — targets the original estimand; no goalpost shift

One polynomial regression step. Same causal parameter. No cost when overlap is fine.


1 ROBUST TO MODEL MISSPECIFICATION
You don't need the true model.
You need the right identification.
2 WORKS IN REALISTIC SAMPLE SIZE
Designed for messy data.
Built to perform.
3 SPANS RESEARCH DESIGNS
IV
DiD
Unconf
One framework.
Many strategies. One goal.
4 RESILIENT TO WEAK COVARIATE OVERLAP
Makes the most of the data you have.
Stays reliable when overlap is weak.
5 STAYS ON TARGET
Estimates the parameter YOU choose,
not one nearby.
6 EASY TO USE
CAUSAL SCIENCE
Simple to implement,
understand, and communicate.
ECONOMETRICS
CAUSAL INFERENCE
STATISTICS
Identification
• Clarity
• Credibility
• Causality
Estimation
• Robust
• Transparent
• Actionable
TRUTH. IDENTIFICATION. IMPACT.
Case Notes
• Define target
• Check identification
• Choose strategy
• Estimate & interpret
POTENTIAL OUTCOMES
IDENTIFICATION
BETTER DECISIONS

Practical Recommendations

1. **Plot the PS distribution** – if no mass near 0 or 1, DR-BC = standard DR
2. **Report trimming fraction** – small → variance reduction; large → polynomial does real work
3. **Check h-sensitivity** – stable across $h \in [0.03, 0.10]$ is reassuring
4. **Default**: $h = 0.05$, $k = 1$, $K = 3$

> There is no reason not to use DR-BC: it adds one polynomial regression, targets the same estimand, and is at worst harmless.

Thank You

pedro.santanna@emory.edu

Backup

Backup: Simulation DGP

- DiD setting: 2 periods, 2 groups, $n = 10{,}000$, 4 Student-$t(10)$ covariates
- Nonlinear transformations induce misspecification in outcome and/or PS model
- True ATT $= 0$; 10,000 MC replications; DR-BC: $h = 0.05$, $k = 1$, $K = 3$

Backup: Rate Conditions on h

- **Upper bound** (bias control): $nh^{2k} = O(1)$
- **Lower bound** (variance control): $nh^4 \to \infty$
- For $k = 1$: both reduce to $h \sim n^{-1/2}$ (degenerate, but feasible for finite n with different constants)
- Fixed $h = 0.05$: residual bias $\leq C \cdot 0.05^2 = 0.0025C$; at $n = 10{,}000$: $n^{-1/2} = 0.01$
- We recommend $k = 1$ because $k \geq 2$ requires sieve derivatives that are imprecisely estimated with fixed $K = 3$

Backup: Formal Assumptions

- **Assumption 1** (Approximate double robustness): If either $m^* = m$ or $p^* = p$, then $\theta_h(\gamma^*) = \theta_0 + o(h^k)$.
- **Assumption 2** (Conditional mean at zero): For each $\ell = 1, \ldots, L$ with $0 \in \textbf{support}(A_\ell(\gamma^*))$:
  (i) $\xi_\ell(0; \gamma^*) = 0$ if $p^* = p$, and $\xi_\ell(a; \gamma^*) = 0$ for all a for the residual moments if $m^* = m$
  (ii) $\xi_\ell(\cdot; \gamma^*)$ is $(k+1)$-times continuously differentiable near 0
- **Assumption 3** (Smoothness of Λ): $\Lambda(\cdot)$ is twice continuously differentiable near $(\alpha_1(0, \gamma^*), \ldots, \alpha_L(0, \gamma^*))$.
- **Assumption 4** (Regularity conditions): For each ℓ:
  (a) First-stage IF: $\alpha_\ell(h, \hat{\gamma}) - \alpha_\ell(h, \gamma^*) = (\mathbb{E}_n[\cdot] - \mathbb{E}[\cdot])[\phi_\ell] + o_p(n^{-1/2})$
  (b) Sieve: $\hat{\xi}_\ell^{(\kappa)}(0; \gamma^*) - \xi_\ell^{(\kappa)}(0; \gamma^*) - (\mathbb{E}_n[\cdot] - \mathbb{E}[\cdot])[\psi_{\ell,\kappa}] = o_p(n^{-1/2}h^{1-\kappa})$
  (c) Moment bound: $\mathbb{E}[\omega_\ell(h, \gamma^*)^2] = o(n^{1/2})$
  (d) Stochastic equicontinuity
  (e) Rate: $nh^{2k} = O(1)$; complementary lower bound $nh^4 \to \infty$ (Appendix)

← Go back

Backup: Full Influence Function

The IF of each moment $\hat{\alpha}_\ell$ has four components:

$$\omega_\ell(h,\gamma) = \underbrace{\frac{B_\ell(\gamma)}{A_\ell(\gamma)}\mathbb{1}\{|A_\ell(\gamma)| \geq h\}}_{\text{trimmed ratio}} + \underbrace{\sum_{\kappa=1}^{k} \frac{A_\ell(\gamma)^{\kappa-1}\mathbb{1}\{|A_\ell(\gamma)| < h\}}{\kappa!} \cdot \xi_\ell^{(\kappa)}(0;\gamma)}_{\text{bias correction}}$$

$$+ \underbrace{\sum_{\kappa=1}^{k} \frac{\mathbb{E}\left[A_\ell(\gamma)^{\kappa-1}\mathbb{1}\{|A_\ell(\gamma)| < h\}\right]}{\kappa!} \cdot \psi_{\ell,\kappa}(\gamma)}_{\text{sieve estimation}} + \underbrace{\phi_\ell}_{\text{first stage}}$$

where $\psi_{\ell,\kappa}(\gamma) = q_K^{(\kappa)}(0)' \, \mathbb{E}\left[q_K(A_\ell(\gamma))q_K(A_\ell(\gamma))'\right]^{-1} q_K(A_\ell(\gamma))(B_\ell(\gamma) - \xi_\ell(A_\ell(\gamma);\gamma))$

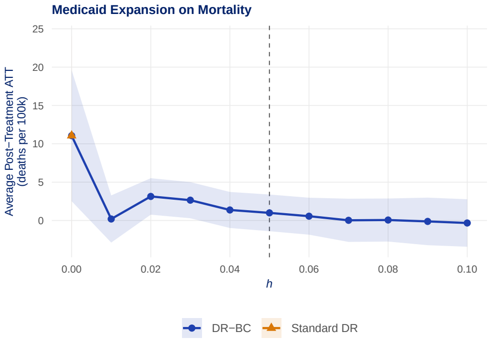
Overall IF via the delta method:

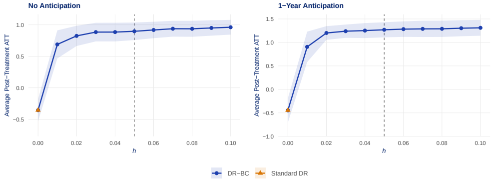
$$\varphi = \sum_{\ell=1}^{L} \Lambda_\ell(\alpha_1(h,\gamma^*),\ldots,\alpha_L(h,\gamma^*))\,\omega_\ell(h,\gamma^*)$$

Variance: $\hat{\sigma}^2 = \mathbb{E}_n\left[(\hat{\varphi} - \mathbb{E}_n[\hat{\varphi}])^2\right]$ — plug sample analogs into ω_ℓ and φ

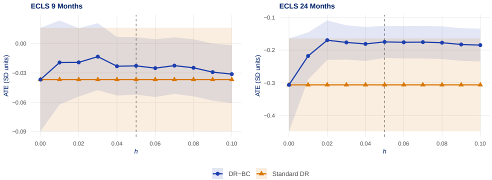
← Go back

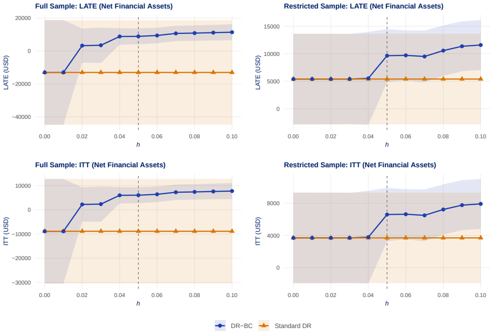
Backup: h-Sensitivity (DiD)





Backup: h-Sensitivity (ATE & LATE)





Backup: 401(k) — How 2 Observations Dominate 9,908

7 obs (0.07% of $n = 9{,}910$) with $\hat{p}_Z(X) > 0.95$ — but AIPW is arm-asymmetric: $Z/\hat{p}_Z(X)$ on eligibles, $(1-Z)/(1-\hat{p}_Z(X))$ on ineligibles.

$\hat{p}_Z(X)$	Z	$1/\hat{p}_Z(X)$	$1/(1-\hat{p}_Z(X))$	Active weight
0.9898	0	1.01	**98**	**98**
0.9676	0	1.03	**31**	**31**
0.9740	1	1.03	38	1.03
0.9731	1	1.03	37	1.03
0.9708	1	1.03	34	1.03
0.9664	1	1.03	30	1.03
0.9556	1	1.05	23	1.05

Two ineligibles carry effective weights 98 and 31 (sample median ≈ 1.5); the other 5 eligibles have benign weights ≈ 1.03. Trimming removes their direct influence; BC puts their contribution back without instability.

Backup: Overlap Weights vs DR-BC

- **Overlap weights** (Li, Morgan and Zaslavsky, 2018): different estimand (overlap-weighted ATE)
- Not extended to LATE or staggered DiD
- DR-BC: targets the **original** estimand (ATE, LATE, or group-time ATT)
- When the research question requires a specific estimand, DR-BC is the natural choice
- When the researcher is willing to change the estimand, overlap weights are a valid alternative

Backup: Limitations & Open Questions

- **Data-driven *h* selection**: currently fixed at $h{=}0.05$; adaptive rule is an open problem
- **Higher-order BC** ($k \geq 2$): finite-sample performance degrades (sieve derivatives noisy with fixed K)
- **Continuous treatments**: current framework is for binary treatment/instrument
- **Very high trimming fractions** ($> 30\%$): polynomial does more work — PS diagnostics essential
- **Efficiency**: semiparametric efficiency bounds under weak overlap remain open

References I

Abadie, Alberto, "Semiparametric instrumental variable estimation of treatment response models," *Journal of econometrics*, 2003, *113* (2), 231–263.

Baker, Andrew, Brantly Callaway, Scott Cunningham, Andrew Goodman-Bacon, and Pedro HC Sant'Anna, "Difference-in-Differences Designs: A Practitioner's Guide," *Journal of Economic Literature*, 2025. Forthcoming.

Belloni, Alexandre, Victor Chernozhukov, Iván Fernández-Val, and Christian Hansen, "Program Evaluation and Causal Inference With High-Dimensional Data," *Econometrica*, 2017, *85* (1), 233–298.

Benjamin, Daniel J., "Does 401(k) eligibility increase saving? Evidence from propensity score subclassification," *Journal of Public Economics*, 2003, *87* (5-6), 1259–1290.

Blandhol, Christine, John Bonney, Magne Mogstad, and Alexander Torgovitsky, "When is TSLS Actually LATE?," *Review of Economic Studies*, 2026. Forthcoming.

Caetano, Carolina and Brantly Callaway, "Difference-in-Differences when Parallel Trends Holds Conditional on Covariates," *arXiv:2406.15288*, 2024.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Chernozhukov, Victor and Christian Hansen, "The impact of 401(k) participation on the wealth distribution: an instrumental quantile regression analysis," *The Review of Economics and Statistics*, 2004, *86* (3), 735–751.

D'Amour, Alexander, Peng Ding, Avi Feller, Lihua Lei, and Jasjeet Sekhon, "Overlap in observational studies with high-dimensional covariates," *Journal of Econometrics*, 2021, *221* (1), 644–654.

Dias, Mateus and Luiz Felipe Fontes, "The Effects of a Large-Scale Mental Health Reform: Evidence from Brazil," *American Economic Journal: Economic Policy*, 2024, *16* (3), 257–289.

Fryer, Roland G. and Steven D. Levitt, "Testing for Racial Differences in the Mental Ability of Young Children," *American Economic Review*, 2013, *103* (2), 981–1005.

Li, Fan, Kari Lock Morgan, and Alan M. Zaslavsky, "Balancing Covariates via Propensity Score Weighting," *Journal of the American Statistical Association*, 2018, *113* (521), 390–400.

Robins, James M, Andrea Rotnitzky, and Lue Ping Zhao, "Estimation of Regression Coefficients When Some Regressors Are Not Always Observed," *Journal of the American Statistical Association*, 1994, *89* (427), 846–866.

References II

Słoczyński, Tymon, “Interpreting OLS Estimands When Treatment Effects Are Heterogeneous: Smaller Groups Get Larger Weights,” *The Review of Economics and Statistics*, 2022, *104* (3), 501–509.

——, “When Should We (Not) Interpret Linear IV Estimands as LATE?,” *Review of Economic Studies*, 2026. Accepted.

Słoczyński, Tymon and Jeffrey M Wooldridge, “A General Double Robustness Result for Estimating Average Treatment Effects,” *Econometric Theory*, 2018, *34* (1), 112–133.

Słoczyński, Tymon, S Derya Uysal, and Jeffrey M Wooldridge, “Abadie’s Kappa and Weighting Estimators of the Local Average Treatment Effect,” *Journal of Business & Economic Statistics*, 2025, *43* (1), 164–177.